A novel Taxonomy for Gestural Interaction Techniques: Considerations for automotive Environments

Adriano Scoditti

A novel taxonomy for gestural interaction techniques: considerations for automotive environments

Adriano Scoditti
Laboratoire d'Informatique de Grenoble, Equipe IIHM
385, rue de la Bibliotheque, BP 53, F-38041 Grenoble cedex 9, France
adriano.scoditti@imag.fr

ABSTRACT

A large variety of gestural interaction techniques is now available. In this article, we use a new taxonomic space [18] as a comparative structure to analyze the applicability of these techniques on automotive environment. The taxonomy plots a gestural interaction technique as a point in a space where the vertical axis denotes the semantic coverage of the technique, and the horizontal axis expresses the physical actions users are engaged in. In addition, syntactic modifiers are used to express the interpretation process of input tokens into semantics, as well as pragmatic modifiers to make explicit the level of indirections between users actions and system responses. In the taxonomy, the complexity of the gestural interaction lexicon, and the syntactic/pragmatic modifiers it is decorated with, are indexes of the cognitive load users are engaged in during the interaction. The integration of modern mobile devices, complex user interfaces and gestural interaction techniques into automotive environment rise the necessity to analyze gestural interaction technique from their cognitive load point of view.

Author Keywords

Handheld devices and mobile computing, Input and interaction technologies, Multi-modal interfaces, Recognition and interpretation of user input (face, body, speech etc.)

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: Miscellaneous

INTRODUCTION

Last generation mobile devices are enhanced with a diversity of sensors capable of probing real world physical properties in real time. The pioneering work on sensor-based interaction techniques [8, 11, 12, 15, 16] has paved the way for an active research area [1, 20, 21]. Although these results satisfy "the gold standard of science" [19], in practice, they are too "narrow truths" [4] to support designers decisions and researchers analysis. Designers and researchers need an overall systematic structure that helps them to reason, compare, elicit (and create!) the appropriate techniques for the problem at hand. Taxonomies, which provide such a structure, are good candidates for generalization in an emerging field. The challenge, however, is to provide a classification framework that is both complete and simple to use. Since completeness is illusory in a moving and prolific domain such as user interface design, we will not include it in our goals.

Figure 1. Integration of last generation mobile devices in automotive environment rise the necessity to analyze gestural interaction technique from their cognitive load point of view [?].

In this article, we propose the interpretation of a new taxonomy for gestural interaction techniques [18] with considerations for automotive environment.

To develop our taxonomy, we have built a controlled vocabulary (i.e. primitives) obtained through an extensive analysis of the taxonomies that have laid the foundations for Human-Computer Interaction (HCI) more than twenty five years ago. For the most part, this early work in HCI has been ignored or forgotten by researchers driven by the trendy "technology push" approach.

Our taxonomy is based on the following principles:

(1) Interaction between a computer system and a human being is conveyed through input (output) expressions that are produced with input (output) devices, and that are compliant with an input (output) interaction language.

(2) As any language, an input (output) interaction language can be defined formally in terms of semantics, syntax, and lexical units.

MIAA 2011, February 13, 2011, Palo Alto, CA, USA.

Figure 2. The "sliding" gesture is semantically multiplexed to achieve different meanings, depending on context.

(3) The generation of an input (output) expression involves using devices whose characteristics, from the human perspective, have a strong impact on the expressiveness and the effectiveness of the user interface [5].

Building on Foley's work [9] as well as on Buxton's pragmatics considerations of input structures [5], our taxonomy brings together the four aspects of interaction ranging from semantics to pragmatics with the appropriate human-motivated extensions for addressing the specificity of gestural interaction based on accelerometers. In contrast to Mackinlay *et al.*'s semantic analysis of the design space for input devices [13], we do not consider the transformation functions that characterize the system-oriented perspective of interaction techniques.

Our expectation is to provide new insights and to start promising directions for the design of novel and powerful gestural interaction techniques.

A NEW TAXONOMY

As shown in Figure 2, the same gesture may convey very different meanings depending on the context in which it is produced: "go to previous photo" as for the Apple's photo album (or "go to next slide" as in Charade in [2]), "open a submenu" in Francone's Wavelet Menu [10], or "unlock" the iPhone screen. In addition, a gesture that makes sense for the system, may not be acceptable in a public social context [17] as it could be meaningful and interpreted by the public itself.

These observations lead us to define a new taxonomy according to the following principles: (1) Coverage of semantic, syntactic, lexical, and pragmatic issues of interaction where semantic granularity is that of Foley's et al. interaction tasks; (2) Adoption of a user centered perspective where physical human actions are premium, leaving aside the internal computational transformations; (3) Consideration for context; (4) Coverage of both foreground and background interaction (as defined by Buxton [6]). Figure 3 shows the elements of the framework that we describe in detail next.

Lexical Axis

Because of our focus on users' involvement in the interaction, the input lexicon corresponds to the physical actions users apply to devices. We divide human physical actions into two groups: (1) conscious actions that belong to the foreground interaction, and (2) unconscious actions that correspond to background interaction. The foreground interaction area contains the interaction techniques that require the user to consciously manipulate the device to reach some objective (as for the sliding gesture of Figure 2). The background interaction area corresponds to the interaction techniques where the system interprets user's unconscious actions together with contextual information to perform some system state change on behalf of the user. For example, during a phone call, the iPhone switches the screen backlight off to safe battery life as the user brings the device next to the ear.

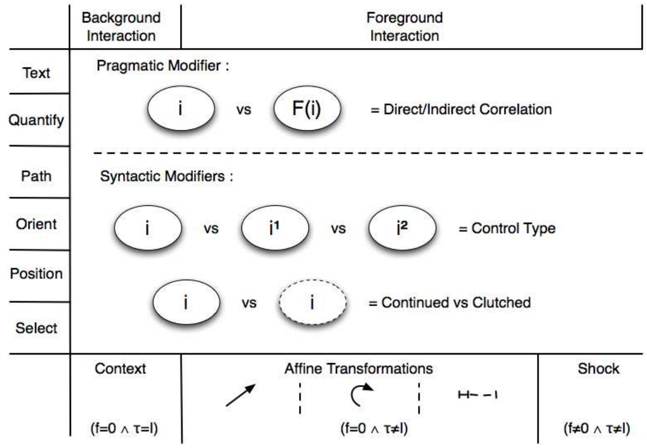

Figure 3. Our classification space for gestural interaction techniques based on accelerometers. The abscissa defines the lexicon in terms of the physical manipulations users perform with the device, with a clear separation between background and foreground interaction. The ordinate corresponds to Foley's interaction tasks. An interaction technique is uniquely identified by an integer *i* and plotted as a point in this space. Each point is decorated with the pragmatic and syntactic properties of the corresponding interaction technique.

Whether human actions are performed consciously to address the system or not, our classification space characterizes these actions with two additional variables: (τ) the geometrical transformation matrix that models user's movements in space, and (f) the frequency of these movements. The combinations of τ and f identify three sub-areas within the lexical axis: "Context", "Affine Transformations" and "Shock". The affine transformations group identifies the most common interaction techniques based on translations, rotations and/or scales (in this case, τ is different from the identity matrix I), and without any repetition (that is, f is equal to zero, meaning that the interaction is time driven). The sliding gesture of Figure 2 falls in this category. The shock category identifies those interaction techniques based on a combination of translations, rotations and/or scales (τ is different from the identity matrix) repeated over time (then, f is different from zero). The shake gesture exemplified by Shoogle [20] falls in this category. The context category corresponds to unconscious human manipulations that the system may interpret to feed into its own context model and, depending on this context, acts on behalf of the user. For this situation, we stipulate that τ is the Identity matrix and f is equal to zero.

Syntactic Axis

Independently from the device used, we characterize the syntactic dimension of an interaction technique with the following two variables that we call *syntactic modifiers*: (1) the existence (or absence) of triggers to specify the begin/end of the interaction, and (2) the control type associated with the input token, which may be position-control, speed-control or acceleration-control. As a result, given that, in our taxonomy, an interaction technique is uniquely identified by an index *i*, the trigger syntactic modifier is represented as an oval that surrounds the interaction technique identifier using a dashed-line or a continuous line to respectively denote the presence (i.e. clutch) or absence (i.e unclutch) of a trigger. In addition, a derivative-like notation is used to convey the control type where *i* is decorated with an exponential number that expresses the derivative order with respect to time (i.e., no derivative for position, first order derivative for speed, and second order derivative for acceleration).

Semantic Axis

As justified in our review about the foundational taxonomies developed in HCI, we re-use Foley's interaction tasks: Select, Position, Orient, Path, Quantify, and Text [9] (See the vertical axis of Figure 3).

Pragmatic Axis

One of the originalities of our work is the attempt to classify gestural interaction techniques in close connection with their meaning in the user's real world. To do this, we introduce a *pragmatic modifier* that expresses the directness [14, 3] of the mapping between the user's expectation (i.e. goal) and the semantics of the interaction technique in the computer world. For indirect mapping, the identifier *i* of the interaction technique becomes the parameter of a function *F(i)* to indicate the existence of one or several reinterpretation layers, whereas for direct mapping, *i* does not receive any additional decoration.

DISCUSSION AND RESEARCH DIRECTIONS

Our fine-structured, language-inspired analysis allows to understand intrinsic and implicit differences even among apparently similar interaction techniques allowing researcher to better explore them and designers to better choose the best suitable for each case.

From the researcher's point of view, the classification shows a transparent state of the art where each interaction technique is classified without ambiguity. Typically, reference taxonomies such as [9] or [5] do not consider the role of time (cf. frequency and duration), nor do they cover unconscious interaction (cf. background interaction) and unstructured interaction such as device shaking. In addition, they do not explicitly consider whether an interaction technique is clutched or unclutched introducing ambiguities and mixing up different aspects of human interaction behavior.

From the designer's point of view, the dimensions of our taxonomy can be used as a framework for decision making. For example, an unclutched interaction technique may be considered for default tasks, while different clutched interaction techniques can be multiplexed through the use of standard or ad-hoc widgets. By proposing at least an interaction technique for each of the proposed task while designing an application, designers will be able to offer a complete and uniform user experience similar to the WIMP one. Furthermore, designers can predict the difficulties that final users will encounter by analyzing the pragmatic and syntactic modifiers that characterize the interaction techniques they envision. Thus, they will be able to choose interaction techniques that best suit the targeted representative users (novice, intermediate, expert).

We think good research and development directions will be both toward the creation of widgets able to transform direct interactions in their more complex counterparts and toward the definition of the elementary interactions to base the development on. The classification suggests to concentrate the efforts toward the development of interaction techniques able to specify Path, Quantity and Text input.

Direct pragmatical interaction techniques are the most suitable for automotive environment, in particular for drivers. The lack of indirection layers during the interaction characterizes lower cognitive loads thus easing the interaction and avoiding distraction.

CONCLUSIONS

The characteristics on which we choose to perform our analysis are the ones inspired by the parallelism existing between artificial languages proposed by interactions and gestural languages users are used to: lexicon, syntax, semantic and pragmatic. Our discussion did not deepened to system level, as we didn't want to differentiate interaction techniques by their implementation characteristics (granularity, resolution function, state machine are the variables already been taken into account [7, 13] whom we want to be complementary rather than substitutes).

Our approach proposed a user-centered classification able to analyze the state of the art of accelerometers-based interaction techniques by the manipulation point of view: the user perform a physical action in its space in order to communicate with the system. We think this is the atomic level on which we have to conceive our interfaces in order to propose system-wide coherent languages to the users. This coherence will drive them through a more agreeable, *natural* [5] and *intuitive* system, having coherence and direct pragmatic distances.

We proposed the use of a parametrical space where the pragmatic distance and the syntactical modifiers are indexes of the learning curve users have to go over when approaching a new interaction language.

We contextualized our approach and principles to automotive environment. We proposed the use of the syntactical and pragmatical modifiers as discriminants of the most appropriate gestural interaction techniques suitable in automotive environments.

REMARKS

The content of this article refers to, and in some part is an extract of, the accelerometers interaction techniques taxonomy proposed by Scoditti *et al.* [18].

REFERENCES

1. R. Ballagas, J. Borchers, M. Rohs, and J. G. Sheridan. The smart phone: A ubiquitous input device. *IEEE Pervasive Computing*, 5(1):70, 2006.
2. T. Baudel and M. Beaudouin-Lafon. Charade: remote control of objects using free-hand gestures. *Commun. ACM*, 36(7):28–35, 1993.
3. M. Beaudouin-Lafon. Instrumental interaction: an interaction model for designing post-wimp user interfaces. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 446–453, New York, NY, USA, 2000. ACM.
4. F. P. Brooks. Grasping reality through illusion—interactive graphics serving science. In *CHI '88: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1–11, New York, NY, USA, 1988. ACM.
5. W. Buxton. Lexical and pragmatic considerations of input structures. *SIGGRAPH Comput. Graph.*, 17(1):31–37, 1983.
6. W. Buxton. Integrating the periphery and context: A new model of telematic. *Proceedings of Graphics Interface*, pages 239–246, 1995.
7. S. K. Card, J. D. Mackinlay, and G. G. Robertson. A morphological analysis of the design space of input devices. *ACM Trans. Inf. Syst.*, 9(2):99–122, 1991.
8. G. W. Fitzmaurice, S. Zhai, and M. H. Chignell. Virtual reality for palmtop computers. *ACM Trans. Inf. Syst.*, 11(3):197–218, 1993.
9. J. D. Foley, V. L. Wallace, and P. Chan. The human factors of computer graphics interaction techniques. *IEEE Comput. Graph. Appl.*, 4(11):13–48, 1984.
10. J. Francone, G. Bailly, L. Nigay, and E. Lecolinet. Wavelet menu: une adaptation des marking menus pour les dispositifs mobiles. In *IHM '09: Proceedings of the 21st International Conference on Association Francophone d'Interaction Homme-Machine*, pages 367–370, New York, NY, USA, 2009. ACM.
11. K. Hinckley, J. Pierce, M. Sinclair, and E. Horvitz. Sensing techniques for mobile interaction. In *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*, pages 91–100, New York, NY, USA, 2000. ACM.
12. G. Levin and P. Yarin. Bringing sketching tools to keychain computers with an acceleration-based interface. In *CHI '99: CHI '99 extended abstracts on Human factors in computing systems*, pages 268–269, New York, NY, USA, 1999. ACM.
13. J. Mackinlay, S. K. Card, and G. G. Robertson. A semantic analysis of the design space of input devices. *Hum.-Comput. Interact.*, 5(2):145–190, 1990.
14. D. Norman. *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc., 1986.
15. K. Partridge, S. Chatterjee, V. Sazawal, G. Borriello, and R. Want. Tilttype: accelerometer-supported text entry for very small devices. In *UIST '02: Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 201–204, New York, NY, USA, 2002. ACM.
16. J. Rekimoto. Tilting operations for small screen interfaces. In *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 167–168, New York, NY, USA, 1996. ACM.
17. J. Rico and S. Brewster. Usable gestures for mobile interfaces: evaluating social acceptability. In *CHI '10: Proceedings of the 28th international conference on Human factors in computing systems*, pages 887–896, New York, NY, USA, 2010. ACM.
18. A. Scoditti, J. Coutaz, and R. Blanch. A novel taxonomy for gestural interaction techniques based on accelerometers. In *IUI 2011*. ACM, 2011.
19. M. Shaw. What makes good research in software engineering? *International Journal of Software Tools for Technology*, 4(1):1–7, 2002.
20. J. Williamson, R. Murray-Smith, and S. Hughes. Shoogle: excitatory multimodal interaction on mobile devices. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 121–124, New York, NY, USA, 2007. ACM.
21. A. Wilson and S. Shafer. Xwand: Ui for intelligent spaces. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 545–552, New York, NY, USA, 2003. ACM.